

# Capacity for Half-Duplex Line Networks with Two Sources

Tobias Lutz

Inst. for Communications Engineering  
Technische Universität München  
Munich, 80290 Germany  
tobi.lutz@tum.de

Gerhard Kramer

Department of Electrical Engineering  
University of Southern California  
Los Angeles, CA 90089 USA  
gkramer@usc.edu

Christoph Hausl

Inst. for Communications Engineering  
Technische Universität München  
Munich, 80290 Germany  
christoph.hausl@tum.de

**Abstract**—The focus is on noise-free half-duplex line networks with two sources where the first node and either the second node or the second-last node in the cascade act as sources. In both cases, we establish the capacity region of rates at which both sources can transmit independent information to a common sink. The achievability scheme presented for the first case is constructive while the achievability scheme for the second case is based on a random coding argument.

## I. INTRODUCTION

Most wireless networks are half-duplex constrained, i.e. the network nodes cannot transmit and receive simultaneously. In order to handle the half-duplex constraint, transmission protocols deterministically split the time of each network node into transmission and reception periods. This approach is easy to realize since nodes do not have to change rapidly between their transmission and reception modes. However, the approach is suboptimal from an information theoretic point of view. It does not take into account that the throughput of each half-duplex node can be increased by allowing it to choose the transmission-reception patterns in dependence of the information to be sent.

This observation goes back to [1] that introduced a binary, deterministic channel model for half-duplex constrained relays and demonstrated, using the example of a three node line network, that larger rates as compared to time-sharing are possible by modulating the operation modes of the relay based on the underlying information [2]. In [3], the capacity of the degraded half-duplex relay channel was derived. The authors also noted that the schedule of the relay has to carry information in order to achieve the capacity.

An extension of this result to line networks with multiple sources was presented in [4] and [5]. In what follows, we refer to the intermediate nodes, i.e., the second to second-last nodes in the cascade, as *relays* since they must relay the first node's message to the last node that is the destination for all messages. Within the setup of [5] a source and a subset of the relays deliver independent information to the destination under the assumption that adjacent node pairs are connected by noise-free  $(q + 1)$ -ary pipes. A coding scheme based on timing was proposed, and based on the asymptotic behavior of the coding scheme, the capacity of deterministic relay cascades of arbitrary length and a single source was

established. If the cascade includes a certain number of relays with their own information, the coding scheme achieves the cut-set bound provided that the rates of the relay sources fall below individual thresholds.

In the present paper, we treat deterministic half-duplex line networks with two sources where either the first or the last relay in the cascade is the second source. In both cases, we establish the capacity region of rates at which both sources can transmit independent information to a common sink.

If the first relay acts as a source, it is shown that the capacity region is the cut-set region. This improves a result derived in [5] which says that the cut-set bound is achievable if the rate of the relay source falls below a certain threshold. In order to understand the new step in the achievability scheme, we briefly describe the scheme in [5]. Therein, the source node encodes its information by means of transmission symbols and idle symbols. An idle symbol indicates a channel use without transmission. The relays encode received information with the transmission pattern and with the value of the transmission symbols. Based on this, codes can be constructed which allow the nodes to cooperate in a sense that each node controls the transmission pattern applied by the next node. Hence, new information injected by the relay source is not allowed to be represented by the transmission pattern since, otherwise, the previous node is not able to control the applied transmission pattern. This is, in fact, the reason why the rate of the relay sources cannot exceed a certain threshold. In the new scheme, the relays still use the original idea. However, the source regards its link to the relay source as an erasure channel where the erasures are a consequence of the half-duplex constraint. Hence, the source and the relay source do not cooperate anymore which enables the relay source to represent own information with transmission patterns. It turns out that this new step is necessary to achieve all points in the cut-set region.

In the second part of this paper, we focus on the case where the last relay in the cascade acts as a source. The capacity region is derived by means of a random coding argument.

## II. NETWORK MODEL

Consider a discrete memoryless relay cascade as depicted in Fig. 1. Each node is labeled by a distinct number from  $\mathcal{V} = \{0, \dots, m\}$  with  $m > 0$ . The integers 0 and  $m$  refer to

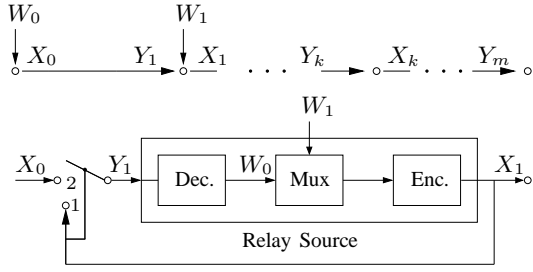


Fig. 1. A noiseless relay cascade with two sources. The link model is illustrated by means of feedback. If relay 1 is transmitting, the switch is in position 1 otherwise in position 2.

the source and sink, respectively, while all remaining integers 1 to  $m - 1$  represent half-duplex constrained relays, i. e. relays which cannot transmit and receive at the same time. The connectivity within the network is described by the set of edges  $\mathcal{E} = \{(k, k + 1) : 0 \leq k \leq m - 1\}$ , i.e. the ordered pair  $(k, k + 1)$  represents the communications link from node  $k$  to node  $k + 1$ . The output of the  $k$ th node, which is the input to channel  $(k, k + 1)$  is denoted as  $X_k$  and takes values on the alphabet  $\mathcal{X}_k = \mathcal{Q}_k \cup \{N\}$  where  $\mathcal{Q}_k$  denotes the transmission alphabet of node  $k$  while the idle symbol “N” signifies a channel use in which node  $k$  is not transmitting. The input of the  $k$ th node, which is the output of channel  $(k - 1, k)$  is denoted as  $Y_k$  and is given by

$$Y_k = \begin{cases} X_{k-1}, & \text{if } X_k = N \\ X_k, & \text{if } X_k \in \mathcal{Q}_k \end{cases} \quad (1)$$

$$Y_m = X_{m-1}. \quad (2)$$

where  $1 \leq k \leq m - 1$ . Channel model (1) captures the half-duplex constraint as follows. Assume relay  $k$  is in its transmission mode, i.e.  $X_k \in \mathcal{Q}_k$ . Then relay  $k$  hears itself ( $Y_k = X_k$ ) but cannot listen to relay  $k - 1$  or, equivalently, relay  $k$  and relay  $k - 1$  are disconnected. However, if relay  $k$  is not transmitting, i. e.  $X_k = N$ , it is able to listen to relay  $k - 1$  via a noise-free  $|\mathcal{X}_{k-1}|$ -ary pipe ( $Y_k = X_{k-1}$ ). Another interpretation of the channel model is that the output  $X_k$  of each relay  $k$  controls the position of a switch which is placed at its input. If relay  $k$  is transmitting, the switch is in position 1 otherwise it is in position 2 (see Fig. 1). Since a pair of nodes is either perfectly connected or disconnected, we obtain a deterministic network with  $p(y_1, \dots, y_m | x_0, \dots, x_{m-1}) \in \{0, 1\}$ .

At the beginning of a new block  $b$  of  $n$  channel uses, source node 0 and relay  $k \in \{1, m - 1\}$  produce a uniformly and independently drawn message  $W_{0,b} \in \{1, \dots, 2^{nR_0}\}$  and  $W_{k,b} \in \{1, \dots, 2^{nR_k}\}$ , respectively. Based on the received sequence in block  $b$ , sink node  $m$  forms the estimates  $\hat{w}_{0,b-(m-1)}$  and  $\hat{w}_{k,b-(m-1-k)}$  of  $W_{0,b-(m-1)}$  and  $W_{k,b-(m-1-k)}$ . We assume the following encoding functions

$$x_{0i} = f_{0i}(W_0) \quad (3)$$

$$x_{ki} = f_{ki}(W_k, Y_k^{i-1}) \quad (4)$$

$$x_{li} = f_{li}(Y_l^{i-1}), \quad \forall l \neq \{0, k, m\}. \quad (5)$$

The first subscript describes the node number while the second subscript  $i$  corresponds to the time instance where  $1 \leq i \leq n$ .

Moreover,  $Y_k^{i-1}$  is used as short hand notation for the set  $\{Y_{k1}, \dots, Y_{k,i-1}\}$ .

### III. THE FIRST RELAY IS A SOURCE

*Theorem 1:* The capacity region  $\mathcal{C}$  of the line network of Fig. 1, where node 0 and relay node 1 are sources, is

$$\mathcal{C} = \bigcup \left\{ \begin{array}{l} R_0 \leq H(Y_1 | X_1) \\ R_0 + R_1 \leq H(Y_m) \\ R_0 + R_1 \leq \min_{2 \leq i \leq m-1} H(Y_i | X_i) \end{array} \right\}. \quad (6)$$

The union is over all probability distributions of the form

$$P_{X_0}(\cdot) P_{X_1}(\cdot) P_{X_2|X_1}(\cdot) P_{X_3|X_2}(\cdot) \dots P_{X_{m-1}|X_{m-2}}(\cdot). \quad (7)$$

*Proof:* We start with the achievability of  $\mathcal{C}$ . At the end of block  $b - 1$ , node 0 and relay node 1 choose new messages  $w_{0,b}$  and  $w_{1,b}$ , respectively, which are sent in block  $b$  by means of the sequences  $\underline{x}_0(w_{0,b})$  and  $\underline{x}_1(w_{0,b-1}, w_{1,b})$ . The remaining relays  $i$ ,  $2 \leq i \leq m - 1$ , forward older messages. In particular, relay  $i$  sends  $\underline{x}_i(w_{0,b-i}, w_{1,b-(i-1)})$  in block  $b$ .

*Coding:*

- At node  $m - 1$  [5]: Node  $m - 1$  represents information by taking  $n_{m-1} < n$  transmission symbols per block of length  $n$  from the alphabet  $\mathcal{Q}_{m-1}$  and by allocating the  $n_{m-1}$  symbols to the transmission block. Thus,  $|\mathcal{Q}_{m-1}|^{n_{m-1}} \binom{n}{n_{m-1}}$  different sequences  $\underline{x}_{m-1}(w_{0,b-(m-1)}, w_{1,b-(m-2)})$  are available at relay  $m - 1$ . Observe that  $|\mathcal{Q}_{m-1}|^{n_{m-1}}$  equals the number of possible distinct sequences when the  $|\mathcal{Q}_{m-1}|$ -ary symbols are located at fixed slots while  $\binom{n}{n_{m-1}}$  equals the number of possible transmission-listen patterns.
- At node  $i$ ,  $1 \leq i \leq m - 2$  [5]: For each transmission-listen pattern used by node  $i + 1$ , node  $i$  generates a codebook. For a particular pattern, node  $i$  allocates  $n_i$  transmission symbols from the alphabet  $\mathcal{Q}_i$  in all possible ways to the  $n - n_{i+1}$  listen slots of the pattern. The slots of the pattern, in which node  $i + 1$  transmits, are filled with idle symbols “N”. This procedure generates a certain number of transmission-listen patterns used by node  $i$ .
- Due to the above codebook construction, adjacent nodes can cooperate since each node  $i \geq 1$  knows the messages to be forwarded by the next node and, thus, is aware of the applied codeword. The construction guarantees that adjacent nodes  $i$  and  $i + 1$ ,  $i \geq 1$ , do not transmit at the same time.
- At node 0: In contrast to [5], node 0 does not adapt to the transmission-listen patterns used by node 1. Instead it uses an optimal point to point erasure channel code with alphabet  $\mathcal{X}_0$  for encoding  $W_{0,b}$ . Output symbols  $Y_1$  of link  $(0, 1)$  are erased with a probability of  $1 - p_{X_1}(N)$ , i.e. the erasure probability is equal to the fraction of time in which node 1 transmits. It should be noted that node 0 transmits a part of the information in the timing of the transmission symbols since the erasure code makes use of symbol “N”.

*Achievable Rates:* The capacity of a  $|\mathcal{X}_0|$ -ary erasure channel with erasure probability  $1 - p_{X_1}(N)$  equals  $p_{X_1}(N) \log |\mathcal{X}_0|$  achieved by a uniform input distribution over  $\mathcal{X}_0$ . Due to the channel model, we clearly have

$$\begin{aligned} H(Y_1|X_1) &= H(X_0|X_1 = N) \\ &\leq p_{X_1}(N) \log |\mathcal{X}_0| \end{aligned} \quad (8)$$

with equality if  $p_{X_0|X_1}(\cdot|N)$  is the uniform distribution over  $|\mathcal{X}_0|$ . Thus, an optimal erasure channel code for the link  $(0, 1)$  satisfies  $R_0 = H(Y_1|X_1) - \epsilon$  with  $\epsilon \rightarrow 0$  as  $n \rightarrow \infty$ .

Further, we know from the results derived in [5, Sec. III.A] that

$$|\mathcal{Q}_i|^{n_i} \binom{n - n_{i+1}}{n_i} \rightarrow 2^{nH(Y_{i+1}|X_{i+1})} \quad \text{as } n \rightarrow \infty. \quad (9)$$

for  $0 \leq i \leq m-2$ . For  $i = m-1$ , the exponent in (9) becomes  $H(Y_m)$  (with  $n_m = 0$ ). Hence,  $R_0 + R_1 \leq \min_{2 \leq i \leq m-1} H(Y_i|X_i)$  and  $R_0 + R_1 \leq H(Y_m)$ .

The converse is immediate since the bounds of (6) correspond to the cut-set upper bound [5, Sec. IV]. (7) follows from the following consideration. Again, due to the channel model

$$H(Y_i|X_i) = p_{X_i}(N)H(X_{i-1}|X_i = N), \quad (10)$$

so that  $H(Y_i|X_i)$  is a function of  $p_{X_{i-1}X_i}(\cdot)$  for all  $2 \leq i \leq m-1$ . Hence, without restriction we may assume the Markov chain  $X_1 - \dots - X_{m-1}$ . Further, we can choose a uniform  $p_{X_0|X_1}(\cdot|N)$  over  $|\mathcal{X}_0|$  since this achieves the upper bound (8). Clearly, such a distribution also exists when  $X_0$  is independent of  $X_1, \dots, X_{m-1}$ . ■

*Remark 1:* Theorem 1 shows that the capacity region of the considered line network is equal to the cut-set region. This improves a result in [5] which says that the cut-set bound is achieved when the rate of the relay source falls below a certain threshold. The new ingredient here is that the relay source is allowed to encode its own information in the timing of transmission symbols. In fact, node 0 accepts that a part of its information is erased by node 1. We point out that this approach, namely to treat the link to the relay source  $k$  as an erasure channel, is not cut-set bound achieving if  $k \geq 2$ .

*Example 1:* We apply Theorem 1 to a line network composed of three nodes where the first two nodes have their own information. The alphabets are  $\mathcal{X}_0 = \mathcal{X}_1 = \{0, 1, N\}$ . This example has already appeared in [4], [5]. However, we are now able to characterize the complete capacity region. Moreover, the approach here is easier since we can restrict attention to independent  $X_0$  and  $X_1$ . By choosing  $P_{X_0}(\cdot)$  to be the uniform distribution and, further, by assigning the same probability masses to  $X_1 = 0$  and  $X_1 = 1$  (due to symmetry), we obtain the following expression for the capacity region

$$\mathcal{C} = \bigcup \left\{ \begin{array}{l} R_0 \leq p_{X_1}(N) \log 3 \\ R_0 + R_1 \leq (1 - p_{X_1}(N)) \log 2 + h(p_{X_1}(N)) \end{array} \right\}. \quad (11)$$

The union is over  $p_{X_1}(N)$  and  $h(\cdot)$  denotes the binary entropy function.  $\mathcal{C}$  is depicted in Fig. 2. Note that the region bounded by the dashed line contains the rates which are achievable

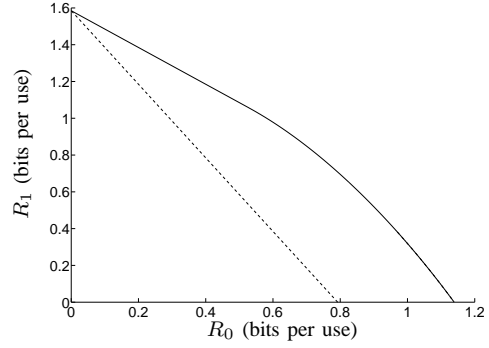


Fig. 2. Capacity region (11) is given by the solid curve. The time-sharing region is bounded by the dashed line.

when the time of the relay is deterministically split into transmission and reception periods. In order to obtain this region, time-sharing between  $(R_0, R_1) = (0.5 \log_2 3, 0)$  and  $(0, \log_2 3)$  bits per use has to be performed.

#### IV. THE LAST RELAY IS A SOURCE

In the remainder, we will make use of the following notation

$$\begin{aligned} w_{0,b-[i+j;i+k]} &\stackrel{\text{def}}{=} \{w_{0,b-(i+j)}, \dots, w_{0,b-(i+k)}\} \\ X_{[l;t]} &\stackrel{\text{def}}{=} \{X_l, \dots, X_t\}. \end{aligned}$$

*Lemma 1:* [6, Th. 14.2.3] Let  $A_\epsilon^{(n)}$  denote the typical set for the probability mass function  $p(x_1, \dots, x_n)$  and let

$$\begin{aligned} P(\underline{X}_1 = \underline{x}_1, \dots, \underline{X}_n = \underline{x}_n) &= \prod_{l=1}^n p(x_{1l}|x_{3l}, \dots, x_{nl}) \\ &\quad p(x_{2l}|x_{3l}, \dots, x_{nl})p(x_{3l}, \dots, x_{nl}). \end{aligned}$$

Then

$$P\{(\underline{X}_1, \dots, \underline{X}_n) \in A_\epsilon^{(n)}\} \doteq 2^{-n(I(X_1; X_2|X_3, \dots, X_n) \pm 6\epsilon)}.$$

*Theorem 2:* The capacity region  $\mathcal{C}$  of the line network of Fig. 1, where node 0 and relay node  $m-1$  are sources, is

$$\mathcal{C} = \bigcup \left\{ \begin{array}{l} R_0 \leq \min_{1 \leq i \leq m-1} H(Y_i|X_i) \\ R_{m-1} \leq H(Y_m|U) \\ R_0 + R_{m-1} \leq H(Y_m) \end{array} \right\}. \quad (12)$$

The union is over all probability distributions of the form

$$P_{X_0} P_{X_1} P_{X_2|X_1} P_{X_3|X_2} \dots P_{X_{m-2}|X_{m-3}} P_{U|X_{m-2}} P_{X_{m-1}|U}. \quad (13)$$

*Remark 2:*  $\mathcal{C}$  is equal to the cut-set region if there exists a probability distribution for each boundary point such that  $X_{m-1}$  is independent of  $U$ . Otherwise,  $\mathcal{C}$  is smaller than the cut-set region.

## V. PROOF OUTLINE OF THEOREM 2

### A. Achievability

*Random codebook generation:*

- Split  $W_0$  into  $B$  sub-blocks  $W_{0,b}$ ,  $b = 1, 2, \dots, B$ , that each take on  $2^{nR_0}$  values. Similarly, split  $W_{m-1}$  into  $B$  sub-blocks  $W_{m-1,b}$ ,  $b = 1, 2, \dots, B$ , that each take on  $2^{nR_{m-1}}$  values.
- Node  $m-1$  generates at random  $2^{nR_0}$  independent sequences of length  $n$ ,  $\underline{u}(w_{0,b-(m-1)})$ ,  $w_{0,b-(m-1)} \in \{1, \dots, 2^{nR_0}\}$ , according to  $\prod_{l=1}^n p(u_l)$ .
- Codebook at node  $m-1$ : On each of the sequences  $\underline{u}(w_{0,b-(m-1)})$ , node  $m-1$  superposes a random codebook with  $2^{nR_{m-1}}$  codewords  $\underline{x}_{m-1}(w_{0,b-(m-1)}, w_{m-1,b})$  using  $\prod_{l=1}^n p(x_{m-1,l}|u_l)$ .
- Codebook at node  $m-2$ : For each  $\underline{u}(w_{0,b-(m-1)})$ , node  $m-2$  generates  $2^{nR_0}$  independent sequences  $\underline{x}_{m-2}(w_{0,b-[m-2;m-1]})$  according to  $\prod_{l=1}^n p(x_{m-2,l}|u_l)$ .
- Codebook at node  $i$ ,  $0 \leq i < m-2$ : For each  $\underline{x}_{i+1}(w_{0,b-[i+1;m-1]})$  node  $i$  generates at random  $2^{nR_0}$  independent sequences  $\underline{x}_i(w_{0,b-[i;m-1]})$  according to  $\prod_{l=1}^n p(x_{i,l}|x_{i+1,l})$ .

*Encoding:* At the beginning of each block  $b$ , node  $i$ ,  $0 \leq i \leq m-2$ , has the estimates<sup>1</sup>  $\hat{w}_{0,b-i-l}$  of  $w_{0,b-i-l}$ ,  $l \geq 0$ . To send the estimate  $\hat{w}_{0,b-i}$ , node  $i$  selects the codeword  $\underline{x}_i(\hat{w}_{0,b-[i;m-1]})$ .

Similarly, at the beginning of block  $b$ , node  $m-1$  has the estimates  $\{\hat{w}_{0,b-(m-1)-l}, \hat{w}_{m-1,b-l}\}$  of  $\{w_{0,b-(m-1)-l}, w_{m-1,b-l}\}$ ,  $l \geq 0$ . To send the pair  $\{\hat{w}_{0,b-(m-1)}, \hat{w}_{m-1,b}\}$ , node  $m-1$  selects the codeword

$$\underline{x}_{m-1}(\hat{w}_{0,b-(m-1)}, \hat{w}_{m-1,b}).$$

Every node  $i$ ,  $1 \leq i \leq m$ , receives the sequence  $\underline{y}_i(b)$  in block  $b$ .

*Decoding:* At the end of block  $b$ , sink node  $m$  performs the following  $\epsilon$ -typicality check in order to determine  $\hat{w}_{0,b-(m-1)}$  and  $\hat{w}_{m-1,b}$ :

$$\left\{ \underline{u}(\hat{w}_{0,b-(m-1)}), \underline{x}_{m-1}(\hat{w}_{0,b-(m-1)}, \hat{w}_{m-1,b}), \underline{y}_m(b) \right\} \in A_{\epsilon}^{(n)}(U, X_{m-1}, Y_m). \quad (14)$$

By Lemma 1, it follows that the error probability of (14) is smaller than

$$2^{-n(I(U, X_{m-1}; Y_m) - 6\epsilon)}. \quad (15)$$

Further, if the estimate  $\hat{w}_{0,b-(m-1)}$  is known at the sink, the error probability regarding the estimate  $\hat{w}_{m-1,b}$  is smaller than

$$2^{-n(I(X_{m-1}; Y_m|U) - 6\epsilon)}. \quad (16)$$

Similarly, at the end of block  $b$ , node  $i$ ,  $1 \leq i \leq m-1$ , performs the following  $\epsilon$ -typicality check in order to determine

<sup>1</sup>The source knows its own messages. However, for simplicity, we will also denote this message with a hat. The same is done for the relay source  $m-1$ .

$\hat{w}_{0,b-i}$ :

$$\left\{ \underline{x}_{i-1}(\hat{w}_{0,b-[i-1;m-1]}), \underline{x}_i(\hat{w}_{0,b-[i;m-1]}), \underline{y}_i(b) \right\} \in A_{\epsilon}^{(n)}(X_{i-1}, X_i, Y_i). \quad (17)$$

According to Lemma 1, the error probability of (17) regarding the estimate  $\hat{w}_{0,b-(i-1)}$  is smaller than

$$2^{-n(I(X_{i-1}; Y_i|X_i) - 6\epsilon)}. \quad (18)$$

Now, by considering all possible error events we obtain from (15), (16) and (18) that

$$\mathcal{R} = \bigcup \left\{ \begin{array}{l} R_0 \leq \min_{1 \leq i \leq m-1} H(Y_i|X_i) \\ R_{m-1} \leq H(Y_m|U) \\ R_0 + R_{m-1} \leq H(Y_m) \end{array} \right\} \quad (19)$$

is an achievable region. Observe that the exponents of the error probabilities can be simplified since  $Y_i$  is a function of  $X_i, X_{i-1}$ .

### B. Converse

Consider the following bounds, where  $P_{b,0}$  and  $P_{b,m-1}$  are the average bit error probabilities when decoding  $W_0$  and  $W_{m-1}$  at the destination node  $m$ . For  $1 \leq l \leq m-1$ , we have

$$nR_0(1 - h(P_{b,0})) \quad (20)$$

$$\stackrel{(a)}{\leq} I(W_0; Y_m^n) \quad (21)$$

$$\leq I(W_0; W_{m-1} Y_l^n Y_m^n) \quad (22)$$

$$\stackrel{(b)}{=} \sum_{i=1}^n I(W_0; Y_{li} Y_{mi} | W_{m-1} Y_l^{i-1} Y_m^{i-1}) \quad (23)$$

$$\stackrel{(c)}{=} \sum_{i=1}^n I(W_0; Y_{li} | W_{m-1} Y_l^{i-1} Y_m^{i-1}) \quad (24)$$

$$\stackrel{(d)}{=} \sum_{i=1}^n I(W_0; Y_{li} | W_{m-1} Y_l^{i-1} Y_m^{i-1} X_l^i) \quad (25)$$

$$\stackrel{(e)}{=} \sum_{i=1}^n H(Y_{li} | W_{m-1} Y_l^{i-1} Y_m^{i-1} X_l^i) \quad (26)$$

$$\stackrel{(f)}{\leq} \sum_{i=1}^n H(Y_{li} | X_{li}) \quad (27)$$

$$\stackrel{(g)}{=} nH(Y_l | X_l, Q) \quad (28)$$

$$\stackrel{(h)}{\leq} nH(Y_l | X_l) \quad (29)$$

where

- (a) follows by Fano's inequality
- (b) follows from the chain rule for mutual information and from the independence of  $W_0$  and  $W_{m-1}$
- (c) follows by Markovity
- (d) follows because  $X_l^i$  is a function of  $Y_l^{i-1}$  for all  $1 \leq l < m-1$  and  $X_{m-1}^i$  is a function of  $Y_{m-1}^{i-1}$  and  $W_{m-1}$
- (e) follows because  $W_0$  determines  $X_0^{i-1}, \dots, X_{m-2}^{i-1}$  what, in turn, determines  $Y_1^{i-1}, \dots, Y_{m-2}^{i-1}$  and  $Y_{m-2}^{i-1}$
- (f) conditioning does not increase entropy

- (g) follows by defining  $Q$  to be a time-sharing random variable with  $Y_l := Y_{lQ}$ ,  $X_l := X_{lQ}$ .
- (h) conditioning does not increase entropy

Further, we have the bounds

$$nR_{m-1}(1 - h(P_{b,m-1})) \quad (30)$$

$$\stackrel{(a)}{\leq} I(W_{m-1}; Y_m^n) \quad (31)$$

$$\leq I(W_{m-1}; W_0 Y_m^n) \quad (32)$$

$$\stackrel{(b)}{=} \sum_{i=1}^n I(W_{m-1}; Y_{mi} | W_0 Y_m^{i-1}) \quad (33)$$

$$\stackrel{(c)}{=} \sum_{i=1}^n I(W_{m-1}; Y_{mi} | W_0 Y_m^{i-1} X_{m-1}^{i-1}) \quad (34)$$

$$\stackrel{(d)}{=} \sum_{i=1}^n I(W_{m-1}; Y_{mi} | W_0 Y_{[1;m]}^{i-1} X_{[0;m-1]}^{i-1}) \quad (35)$$

$$\stackrel{(e)}{=} \sum_{i=1}^n H(Y_{mi} | W_0 Y_{[1;m]}^{i-1} X_{[0;m-1]}^{i-1}) \quad (36)$$

$$\stackrel{(f)}{\leq} \sum_{i=1}^n H(Y_{mi} | V_i) \quad (37)$$

$$\stackrel{(g)}{=} nH(Y_m | U) \quad (38)$$

where

- (a) follows by Fano's inequality
- (b) follows from the chain rule for mutual information and from the independence of  $W_0$  and  $W_{m-1}$
- (c) follows because  $Y_m^{i-1} = X_{m-1}^{i-1}$
- (d) follows because  $W_0$  determines  $Y_1^{i-1}, \dots, Y_{m-1}^{i-1}$
- (e) follows because  $W_0$  and  $W_{m-1}$  determine  $Y_{m,i}$
- (f) follows by defining  $V_i = (X_{m-1}^{i-1}, Y_{m-1}^{i-1})$  and from the fact that conditioning does not increase entropy
- (g) follows by defining  $Q$  to be a time-sharing random variable with  $U := (V_Q, Q)$  and  $Y_m := Y_{mQ}$ .

Concerning the sum-rate, we obtain

$$nR_0(1 - h(P_{b,0})) + nR_{m-1}(1 - h(P_{b,m-1})) \quad (39)$$

$$\stackrel{(a)}{\leq} I(W_0; Y_m^n) + I(W_{m-1}; Y_m^n) \quad (40)$$

$$\leq I(W_0; Y_m^n) + I(W_{m-1}; W_0 Y_m^n) \quad (41)$$

$$\stackrel{(b)}{=} I(W_0 W_{m-1}; Y_m^n) \quad (42)$$

$$= \sum_{i=1}^n I(W_0 W_{m-1}; Y_{mi} | Y_m^{i-1}) \quad (43)$$

$$\stackrel{(c)}{\leq} \sum_{i=1}^n H(Y_{mi}) \quad (44)$$

$$\stackrel{(d)}{=} nH(Y_m | Q) \quad (45)$$

$$\stackrel{(e)}{\leq} nH(Y_m). \quad (46)$$

where

- (a) follows by Fano's inequality

- (b) follows from the independence of  $W_0$  and  $W_{m-1}$
- (c) follows since  $W_0$  and  $W_{m-1}$  determine  $Y_{m,i}$  and from the fact that conditioning does not increase entropy
- (d) follows by defining  $Q$  to be a time-sharing random variable and  $Y_m := Y_{mQ}$
- (e) conditioning does not increase entropy.

It remains to check (13). Observe that  $X_{m-1,i}$  is a function of  $Y_{m-1}^{i-1}$  and  $W_2$ . Since  $X_{0i}, \dots, X_{m-2,i}$  do not depend on  $W_2$ , we have the Markov chain  $X_{0i}, \dots, X_{m-2,i} - X_{m-1}^{i-1} Y_{m-1}^{i-1} - X_{m-1,i}$ . Hence, we have

$$P(u, x_0, \dots, x_{m-1}) \quad (47)$$

$$= P(u)P(x_0, \dots, x_{m-1} | u) \quad (48)$$

$$= P(u)P(x_{0i}, \dots, x_{m-2,i} | i, x_{m-1}^{i-1}, y_{m-1}^{i-1}) \quad (49)$$

$$\cdot P(x_{m-1,i} | i, x_{m-1}^{i-1}, y_{m-1}^{i-1}) \quad (50)$$

$$= P(u)P(x_0, \dots, x_{m-2} | u)P(x_{m-1} | u) \quad (50)$$

which shows that

$$X_0, \dots, X_{m-2} - U - X_{m-1}. \quad (51)$$

Finally, by the explanations in the last section of the proof of Theorem 1 we have the Markov chain

$$X_1 - \dots - X_{m-2} - U - X_{m-1} \quad (52)$$

and the independence of  $X_0$  from  $X_1, \dots, X_{m-1}, U$ . ■

## VI. DISCUSSION

An obvious extension is to allow any relay in the cascade to act as second source. However, the solution for this case turns out to be elusive. Though developing achievable rate regions using superposition random coding is straightforward, proving a converse seems to be more difficult. An intuitive explanation is that having the second source located at the first or the last link offers greater freedom for choosing a coding strategy as compared to the other links. This is related to the fact that the first source does not receive information while the sink node does not send information and, therefore, both nodes are not affected by the half-duplex constraint. One could also think of extending the erasure coding technique outlined in the proof of Theorem 1. In particular, if all nodes before the relay source use independent erasure codes, the relay source would be able to send own information in the timing of transmission symbols. However, it can be shown that this approach does not achieve the cut-set bound and, therefore, a converse is missing again.

## ACKNOWLEDGMENT

T. Lutz and C. Hausl were supported by the European Commission in the framework of the FP7 (contract n. 215252). G. Kramer was supported by NSF Grant CCF-09-05235.

## REFERENCES

- [1] G. Kramer. Models and Theory for Relay Channels with Receive Constraints. In *Proc. 42nd Annual Allerton Conf. Commun., Control, and Computing*, (Monticello, IL), Sept. 29 - Oct. 1 2004.
- [2] G. Kramer. Communication Strategies and Coding for Relaying. *Wireless Communications*, vol. 143 of the IMA Volumes in Mathematics and its Applications:163–175, Springer: New York, 2007.

- [3] S. Vijayakumaran, T. F. Wong, and T. M. Lok. Capacity of the Degraded Half-Duplex Relay Channel. *available on arXiv:0708.2270v1 [cs.IT]*.
- [4] T. Lutz, C. Hausl, and R. Kötter. Coding Strategies for Noise-Free Relay Cascades with Half-Duplex Constraint. In *IEEE International Symposium on Information Theory (ISIT)*, pages 2385–2389, Toronto, Ontario, Canada, July 2008.
- [5] T. Lutz, C. Hausl, and R. Kötter. Bits Through Relay Cascades with Half-Duplex Constraint. *submitted to the IEEE Trans. Inform. Theory, available on arXiv:0906.1599v2 [cs.IT]*.
- [6] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley, Inc., 1991.